PROGRAMOWANIE NIELINIOWE

Maciej Patan

Instytut Sterowania i Systemów Informatycznych
Uniwersytet Zielonogórski

WSTĘP

Zadanie programowania nieliniowego (ZPN)

$$\begin{cases} \min f(x) \\ g_i(x) \leqslant 0, \quad i = 1, \ldots, m_g \\ h_i(x) = 0, \quad i = 1, \ldots, m_h \end{cases}$$

$f(x)$ – funkcja celu
$g_i(x)$ i $h_i(x)$ – funkcje ograniczeń

Twierdzenie Weierstrassa (o istnieniu)

Jeżeli funkcja f jest ciągła, a zbiór rozwiązań dopuszczalnych ograniczony i domknięty, to istnieje przynajmniej jedno minimum globalne funkcji f na X.

Warunki konieczne Kuhna-Tuckera (WKT)

Jeżeli $\hat{x} \in X$ jest minimum lokalnym ZPN i jest punktem regularnym, to istnieją liczby $\hat{\lambda}, i = 1, \ldots, m_g$ oraz liczby $\hat{\mu}, i = 1, \ldots, m_h$ takie, że

$$\begin{cases} \nabla f(\hat{x}) + \sum_{i=1}^{m_g} \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=1}^{m_h} \hat{\mu}_i \nabla h_i(\hat{x}) = 0 \\ \hat{\lambda} g_i(\hat{x}) = 0, \quad i = 1, \ldots, m_g \\ \hat{\lambda} \geqslant 0, \quad i = 1, \ldots, m_g \end{cases} \tag{1}$$

Liczby $\hat{\lambda}_i$ oraz $\hat{\mu}_i$ noszą nazwę *mnożników Lagrange'a*

Warunki regularności

Warunek Karlina. Jeżeli funkcje $g_i, i = 1, \ldots, m_g$ oraz $h_i, i = 1, \ldots, m_h$ są liniowe, to każdy punkt $x \in X$ jest regularny

Warunek Slatera. Jeżeli funkcje $g_i, i = 1, \ldots, m_g$ są wypukłe, funkcje $h_i, i = 1, \ldots, m_h$ są liniowe oraz istnieje punkt x^0 taki, że $g_i(x^0) < 0, i = 1, \ldots, m_g$ i $h_i(x^0) = 0, i = 1, \ldots, m_h$ to każdy $x \in X$ jest regularny

Warunek liniowej niezależności. Jeżeli w punkcie x gradienty $\nabla g_i(x), i \in A$ oraz $\nabla h_i(x), i = 1, \ldots, m_h$, tworzą układ wektorów liniowo niezależnych, to punkt x jest regularny

Warunki wystarczające Kuhna-Tuckera

Jeżeli funkcja f jest wypukła, funkcje $g_i, i = 1, \ldots, m_g$ są wypukłe, a funkcje $h_i, i = 1, \ldots, m_h$ są liniowe, to każdy punkt spełniający WKT jest minimum globalnym

Funkcja Lagrange'a (Lagranżian)

Funkcją Lagrange'a związaną z ZPN nazywamy funkcję $L : \mathbb{R}^n \times \mathbb{R}^{m_g} \times \mathbb{R}^{m_h} \times \mathbb{R}^1$ określoną następująco:

$$L(x, \hat{\lambda}, \hat{\mu}) = f(x) + \sum_{i=1}^{m_g} \hat{\lambda}_i g_i(x) + \sum_{i=1}^{m_h} \hat{\mu}_i h_i(x) \tag{2}$$

Za pomocą funkcji Lagrange'a warunki Kuhna-Tuckera można zapisać następująco: istnieją wektory $\hat{\lambda} \in \mathbb{R}^{m_g}$, $\hat{\mu} \in \mathbb{R}^{m_h}$ takie, że

$$\begin{cases} \nabla_x L(\hat{x}, \hat{\lambda}, \hat{\mu}) = 0 \\ \hat{\lambda}^T \nabla_\lambda L(\hat{x}, \hat{\lambda}, \hat{\mu}) = 0 \\ \hat{\lambda} \geqslant 0 \end{cases} \tag{3}$$

PROGRAMOWANIE KWADRATOWE

Sformułowanie problemu

- Funkcja celu

$$z = \frac{1}{2}x^T\mathbf{C}x - \mathbf{d}^T x \to \min$$

x – wektor n-wymiarowy, d – wektor n-wymiarowy, $\mathbf{C}$ – macierz dodatnio określona o rozmiarach $n \times n$

- Ograniczenia

$$\mathbf{A}x \leqslant \mathbf{b}$$

$\mathbf{A}$ – macierz ograniczeń o wymiarach $m \times n$

Właściwości

- funkcja celu jest funkcją kwadratową – różnica formy kwadratowej i formy liniowej
- macierz $\mathbf{C}$ – dodatnio określona
- ze względu na postać funkcji celu jest to problem programowania wypukłego

Przykłady zastosowań

- wyznaczanie zysku ze sprzedaży
- wyznaczanie awaryjnego poziomu zapasów
- wybór akcji do portfela

Rozwiązywanie zagadnień programowania kwadratowego

- Zbudowanie funkcji Lagrange'a
- Utworzenie warunków Kuhna-Tuckera
- Rozwiązanie zadania zastępczego - metoda Wolfe'a

Postępowanie

Zauważmy, że funkcje $h_i = 0, i = 1, \ldots, m_h$

stąd funkcja Lagrange'a

$$L(x, \hat{\lambda}) = f(x) + \sum_{i=1}^{m_g} \hat{\lambda}_i g_i(x) \tag{4}$$

$$L(x, \hat{\lambda}) = \frac{1}{2} x^T \mathbf{C} x - \mathbf{d}^T x + \hat{\lambda}(\mathbf{A}x - \mathbf{b}) \tag{5}$$

warunki Kuhna-Tuckera

$$\begin{cases} \nabla_x L(\hat{x}, \hat{\lambda}) = 0 \\ \hat{\lambda}^T \nabla_\lambda L(\hat{x}, \hat{\lambda}) = 0 \\ \hat{\lambda} \geqslant 0 \end{cases} \qquad \begin{cases} \mathbf{C}x - \mathbf{d}^T + \lambda\mathbf{A} = 0 \\ \lambda^T(\mathbf{A}x - \mathbf{b}) = 0 \\ \hat{\lambda} \geqslant 0 \end{cases} \qquad \begin{cases} \mathbf{C}x + \lambda\mathbf{A} = \mathbf{d}^T \\ \mathbf{A}x = \mathbf{b} \\ \hat{\lambda} \geqslant 0 \end{cases}$$

postać macierzowa

$$\begin{bmatrix} \mathbf{C} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \end{bmatrix} \tag{6}$$

Metoda Wolfe'a

- Największą korzyścią metody jest możliwość zastosowania metody sympleks do rozwiązania zagadniania programowania kwadratowego
- Pierwszym krokiem procedury jest określenie bazowego rozwiązania układu $\mathbf{A}x = \mathbf{b}$ analogicznie jak w metodzie sympleksowej
- Do pierwszego równania w (6) dodaje się nieujemne zmienne sztuczne tak, aby móc określić początkowe rozwiązanie bazowe
- Następnie minimalizuje się sumę tych zmiennych sztucznych, aż równać się ona będzie zeru – szukane rozwiązanie bazowe

Postać zadania pomocniczego

znaleźć minimum

$$z_F = \sum_j u_j$$

przy ograniczeniach

$$\mathbf{Q}\mathbf{w} = \mathbf{f}$$

$$w \geqslant 0$$

gdzie

$$\mathbf{Q} = \begin{bmatrix} \mathbf{C} & \mathbf{A}^T & -\mathbf{A}^T & \mathbf{E} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \mathbf{f} = [\mathbf{d}\ \mathbf{b}]^T, \mathbf{w} = [\mathbf{x}\ \lambda_1\ \lambda_2 \mathbf{u}]$$

$$\hat{\lambda} = \lambda_1 - \lambda_2, \lambda_1 \geqslant 0, \lambda_2 \geqslant 0$$

Przykład. Konstrukcja portfela akcji – model Markowitza

- Minimalizacja ryzyka przy założonym z góry poziomie oczekiwanego zysku
- Kursy akcji cechują się pewną inercją – na podstawie przeszłości można przewidywać przyszłość
- *Oczekiwana stopa zysku* – średnia stopa zysku wyznaczona na podstawie pewnej liczby okresów z przeszłości
- *Miara ryzyka* – odchylenie standardowe lub wariancja stopy zysku
- *Oczekiwany zysk* – funkcja postaci

$$R_p = \sum_{i=1}^{n} x(i) R_i$$

gdzie n – liczba rozpatrywanych spółek, R_p – oczekiwany zysk z portfela, x_i – udział wartościowy akcji i-tej spółki w portfelu, R_i – oczekiwany zysk z akcji i-tej spółki

- *Ryzyko* – funkcja postaci

$$V_p = \sum_{i=1}^{n} \sum_{j=1}^{n} x(i)x(j)S_iS_j\rho_{ij}$$

gdzie V_p – wariancja stopy zysku z portfela, S_i –odchylenie standardowe stóp zysku akcji i-tej spółki, ρ_{ij} współczynnik korelacji stóp zysku akcji i-tej i j-tej spółki

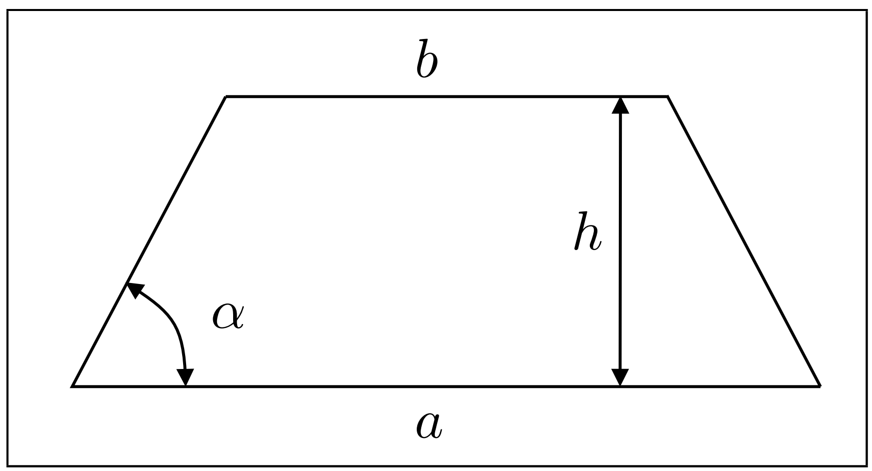
- Portfel optymalny – portfel charakteryzujący się najmniejszym ryzykiem przy założonym poziomie oczekiwanego zysku R_0
- Suma udziałów wszyskich spółek w portfelu musi być równa 1

- W rezultacie otrzymujemy:

$$\begin{aligned} \min\,[V_p] \\ R_p \geqslant R_0 \\ \sum_{i=1}^{n} x(i) = 1 \\ x(i) \geqslant 0, \quad i = 1, \ldots, n \end{aligned}$$

OGRANICZENIA RÓWNOŚCIOWE

Przykład. Metalowa belka ma przekrój w kształcie trapezu. Pole przekroju belki wynika z wymagań wytrzymałościowych konstrukcji i ma wysokość S. Górną i boczne powierzchnie belki pokrywa się kosztownym materiałem antykorozyjnym. Należy określić parametry przekroju tak, aby koszt zabezpieczenia antykorozyjnego był minimalny



- Przekrój optymalny – suma dwóch boków i górnej podstawy ma wartość minimalną
- Funkcja celu

$$\varphi(a, b, h) = b + \sqrt{4h^2 + (a-b)^2}$$

- Ograniczenie

$$\frac{1}{2}h(a+b) - S = 0$$

- Zamiana zmiennych: $x = a + b$, $y = a - b$
- Otrzymujemy:

$$\min \left[f(x, y, h) = \frac{1}{2}(x - y) + \sqrt{4h^2 + y^2} \right]$$

$$g(x, y, h) = S - \frac{1}{2}hx = 0$$

- Poszukujemy minimum globalnego w obszarze $x > 0$, $h > 0$
- W obszarze poszukiwań spełniny jest warunek regularności, gdyż $\nabla g(x, y, h) \neq 0$
- Warunki Kuhna-Tuckera

$$\frac{1}{2} - \frac{1}{2}h\mu = 0, \qquad -\frac{1}{2} + \frac{y}{\sqrt{4h^2 + y2}} = 0, \qquad \frac{4h}{\sqrt{4h^2 + y^2}} - \frac{1}{2}x\mu = 0$$

- Po rozwiazaniu otrzymujemy

$$h = \frac{1}{\mu}, \quad y^2 = \frac{4}{3\mu^2}, \quad x = \frac{2\sqrt{3}}{\mu}$$

- Z postaci funkcji celu wynika, że wartość optymalna y powinna być dodatnia, więc

$$y = \frac{2}{\mu\sqrt{3}}$$

- Mnożnik Lagrange'a μ wynika z ograniczenia równościowego

$$\mu = \frac{\sqrt[4]{3}}{\sqrt{S}}$$

- Ostatecznie otrzymujemy

$$a = \frac{4\sqrt{S}}{\sqrt[4]{27}}, \quad b = \frac{2\sqrt{S}}{\sqrt[4]{27}}, \quad h = \frac{\sqrt{S}}{\sqrt[4]{3}}$$

- Można zauważyć, że

$$\frac{b}{a} = \frac{1}{2}, \quad \frac{h}{a} = \frac{\sqrt{3}}{4}, \quad \alpha = \frac{\pi}{3}$$

- Parametry przekroju nie zależą od S

OGRANICZENIA NIERÓWNOŚCIOWE

Przykład. Rozwiązać zadanie programowania nieliniowego

$$\min \left[f(x) = -(x_1 - 2)^2 - (x_2 - 4)^2 \right] \tag{7}$$

$$g_1(x) = x_1 + x_2 - 8 \leqslant 0 \tag{8}$$

$$x_1 \geqslant 0, \quad x_2 \geqslant 0 \tag{9}$$

- Zbiór dopuszczalny jest domknięty i ograniczony, a funkcja f ciągła zatem na mocy tw. Weierstrassa zadanie ma rozwiązanie
- Spełniony jest warunek regularności Slatera (ograniczenia mają postać liniową), więc rozwiazanie spełnia warunki konieczne Kuhna-Tuckera

- wprowadzamy funkcje $g_2(x) = -x_1$ i $g_3(x) = -x_2$
- tworzymy funkcję Lagrange'a

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \lambda_3 g_3(x)$$

- warunki Kuhna-Tuckera

$$\begin{cases} -2x_1 + 4 + \lambda_1 + \lambda_2 = 0 \\ -2x_2 + 8 + \lambda_1 + \lambda_3 = 0 \\ \lambda_1(x_1 + x_2 - 8) = 0 \\ \lambda_2(-x_1) = 0 \\ \lambda_3(-x_2) = 0 \end{cases}$$

- zadanie ma trzy rozwiązania $A = (0, 8)$, $B = (8, 0)$, $C = (3, 5)$
- dla $A = (8, 0)$ funkcja przyjmuje wartość najmniejszą – rozwiązanie zadania